MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963-A

SECURITY CL_____ (When Data Entered)

AD-A192 717

| REPORT DOCUMENTATION PAGE | READ INSTRUCTIONS BEFORE COMPLETING FORM |
|---|---|

| 1. REPORT NUMBER | 2. GOVT ACCESSION NO. | 3. RECIPIENT'S CATALOG NUMBER |
|---|---|---|
| **AFOSR·TR· 88 - 0 4 3 2** | | |

| 4. TITLE (and Subtitle) | 5. TYPE OF REPORT & PERIOD COVERED |
|---|---|
| Neural Network Research: A Personal Perspective | *Reprint* ~~Interim report~~ |
| | 6. PERFORMING ORG. REPORT NUMBER |

| 7. AUTHOR(s) | 8. CONTRACT OR GRANT NUMBER(s) |
|---|---|
| Stephen Grossberg | ~~AFOSR~~ F49620-86-C-0037 ~~AFOSR~~ F49620-87-C-0018 |

| 9. PERFORMING ORGANIZATION NAME AND ADDRESS | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS |
|---|---|
| Center for Adaptive Systems Department of Mathematics, Boston University Boston, MA 02215 | 61102F 2304/A7 |

| 11. CONTROLLING OFFICE NAME AND ADDRESS | 12. REPORT DATE |
|---|---|
| AFOSR Life Sciences Directorate Bolling Air Force Base Washington, DC 20332 | March 1988 |
| | 13. NUMBER OF PAGES |
| | 9 pages |

| 14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) | 15. SECURITY CLASS. (of this report) |
|---|---|
| *Same as 11* | unclassified |
| | 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE |

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release, distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

Invited article to appear in Electronic Engineering Times

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

These vision preprocessor and ART autonomous classifier examples are just two of the many neural network architectures now being developed by engineers and scientists worldwide. Some of them provide a fertile ground for gaining a new understanding of biological intelligence. Others suggest novel computational theories with natural realizations as real-time adaptive neural network architectures with promising properties for tackling some of the outstanding problems in computer science and technology today. Still others do both. Whatever the focus, here is a field ready to challenge and reward the sustained efforts of a wide variety of gifted people.

DD FORM 1473 EDITION OF...

AFOSR-TR- 88-0432

# NEURAL NETWORK RESEARCH:
# A PERSONAL PERSPECTIVE

Stephen Grossberg†
Center for Adaptive Systems
Boston University
111 Cummington Street
Boston, MA 02215

Invited article to appear in *Electronic Engineering Times*

(An abridged unauthorized version was published in
the March 7, 1988 issue of the *EE Times*)

| Accession For | |
| --- | --- |
| NTIS GRA&I | X |
| DTIC TAB | ☐ |
| Unannounced | ☐ |
| Justification | |
| By | |
| Distribution/ | |
| Availability Codes | |
| Dist | Avail and/or Special |
| A-1 | |

88 5 02 138

# THE LURE OF NEURAL NETWORK RESEARCH

Thousands of scientists, engineers, and students are now studying, developing, or applying neural network models to a wide variety of problems concerning both biological models of brain and behavior and technological models for implementation in government and industrial applications. Many engineers have been drawn to the field because neural network researchers have discovered promising approaches to the many types of problems for which adaptive, massively parallel, fault tolerant solutions are needed, and for which neural networks will run in real-time when they are realized compactly in specialized hardware. In addition, the most advanced neural network architectures are providing examples of intelligent systems capable of autonomous learning and skillful performance within complex and noisy environments that are not under strict control. Such examples and future possibilities have helped to generate a high level of enthusiasm among people working in the field.

No small amount of this enthusiasm also derives from the fact that many neural network design principles, mechanisms, and architectures were discovered through analyses of the human mind and its neural mechanisms. After all, we all have one! What engineer or scientist would not be tempted by the opportunity to understand better this most personal possession, the crucible of all our experience, and even get paid to do so?

The very allure of this possibility goes hand-in-hand with the difficulty, and the greatness, of the task. The human brain is one of the most complex systems amenable to study by human scientists. Not surprisingly, the field of neural network research is remarkable for the diversity and complexity of its subject matter, methods, dialects, and goals.

## DISTINCT CRITERIA FOR BIOLOGICAL AND TECHNOLOGICAL MODELS

In order to find one's bearings and to maintain a high level of sustained productivity in such a challenging field, it is necessary to carefully distinguish between the very different criteria that are appropriate for evaluating biological and technological neural network models. Let me emphasize right away that there is no necessary reason why all studies of biological intelligence should go technological, or conversely. In fact, an uncritical mixing of the two goals can lead to intellectual bankruptcy.

To validly advance our understanding of biological intelligence, one must explain and predict lots of biological data. Typically, there exists one theory in each area of physical science at any time that explains and predicts more data in a principled and harmonious way than any other theory. This becomes the theory worthy of further development, modification, or disconfirmation. Less predictive theories, or theories that are simplified versions or parts of a deeper theory, do not represent the cutting edge of the science, and basically are a waste of time.

Although this seems obvious enough, it is not the way that much Artificial Intelligence has been practiced, and it is not the way that some neural network research has begun to be practiced. For example, the popular back propagation model contains a non-local transport of associative learning weights that is not neurobiologically possible; yet back propagation is starting to be applied by some scientists as a model of neural processes. This is extraordinary, because back propagation was not developed from the study of psychological or neural data. One might just as well try explaining data about electrons using a theory that was developed without regard to how electrons behave. Such a theory could, at best, explain those relatively uninteresting properties of electrons that are model-independent, and, hence, could also be explained, in principle, by a very wide class of alternative models. Such a misapplication of a theory ultimately weakens its case for acceptance, rather than strengthening it, because the theory must then be justified more by expert marketing and hyperbole than by its intrinsic merits.

To advance our understanding of machine intelligence, one need never mention brain or behavior. But one cannot fail to solve outstanding technological problems. In the technological domain, a non-biological neural network model like back propagation can shine, without in the least apologizing for its irrelevance as a brain theory. Moreover, there need not be one best model for each class of technological problems at any time. Multiple models can coexist, just as multiple types of automobiles can all simultaneously be popular in the marketplace.

Diminishing returns set in when a flimsy technological advance is propped up by saying it works just like the brain, or when a metaphorical brain theory devoid of data implications is heralded as the next hi-tech sensation. Then the hyperbole buries the science, much as some AI practitioners have artfully and profitably done, to the ultimate detriment of their field.

## THE INTERNATIONAL NEURAL NETWORK SOCIETY

One good way to clearly distinguish, but still encourage, the very different standards of biological modelling and of technological modelling in the neural network field is to create institutions where both types of activity are equally welcome. From this perspective, the incorporation on March 16, 1987 of the International Neural Network Society (INNS) may eventually prove to be an event of enduring historic importance. Indeed, the seventeen individuals who came together on June 20, 1987 for the first meeting of the INNS governing board included distinguished psychologists, neurobiologists, computer scientists, mathematicians, physicists, and engineers. All of these individuals shared the bond that they are actively engaged in neural network research. Although each of these individuals already belonged to established societies that represent one or more of these disciplines, the full range of neural network research had not previously been supported by any one established society. The INNS was established to provide a bridge between these established societies, and to create a new framework capable of supporting, without confusing their standards, the full range of biological and technological modelling activitites, interdisciplinary knowledge transfers, and cooperative programs that are needed to achieve the great potential of the neural network field. A remarkable and heartwarming measure that INNS fills a much needed niche is that there are already more than 1300 INNS members, with new members joining every day, despite the fact that members began joining only seven months ago.

The two primary vehicles whereby the fledgling INNS now tangibly supports neural network research are its official journal *Neural Networks*, whose first issue came out in January (call 202-767-1493 for further information), and the INNS annual meeting, whose first occurrence, welcoming all facets of neural network research from biology through technology, will take place at the Park Plaza Hotel in Boston on September 6–10, 1988.

A promising sign that the Boston INNS meeting will be a remarkable event is that many established societies which have become interested in neural network research have generously agreed to cooperate with INNS in organizing the meeting. Representing societies whose membership includes a large number of engineers, computer scientists, and physicists are the Computer Society of the IEEE, the IEEE Control Systems Society, the IEEE Engineering in Medicine and Biology Society, the IEEE Systems, Man and Cybernetics Society, the Optical Society of America, and the Society of Photo-Optical Instrumentation Engineers. Mathematicians are represented through the American Mathematical Society and the Society for Industrial and Applied Mathematics. Psychologists and biologists are represented through societies such as the Association for Behavior Analysis, the Cognitive Science Society, the Society for Mathematical Biology, and the Society for the Experimental Analysis of Behavior. This diversity is also reflected in the extraordinary range of engineers and scientists who will be offering tutorials, symposium and plenary lectures, contributed oral and poster presentations, and exhibits (call 617-237-7931 for information

about contributing a paper, exhibiting, and registration).

## THE ARCHITECTURE IS THE ALGORITHM

How can one intellectually distinguish neural network research from other approaches to the study of intelligence and control? To fully clarify this issue, one needs to distinguish several types of neural network contributions, their mutual relationships, and their relationships to other research areas. Although full documentation of these distinctions cannot be developed here, some key points can be made.

First, there exists a continuum of neural network models, from contributions that compete with a number of alternative research approaches, to contributions for which neural nets have offered unique approaches for which no competitors yet exist. The latter type of model represents the truly revolutionary potential of neural network research. But even in the former type, neural network researchers have contributed new computational theories and design ideas to the solution of their targeted problems.

In all of these cases, the neural network model represents a natural realization of the new computational theory. Some people like to summarize this by saying that "the architecture is the algorithm." This direct relationship between these new computational theories and their architectural realizations provides a blueprint to engineers for implementing neural network models in compact, real-time hardware.

## MECHANISMS AND MODULES

One way to group neural network contributions is in terms of mechanisms, design principles, modules, and architectures. Mechanisms include nonlinear differential equations for fast distributed information processing, also called short term memory (STM) equations or activation equations; differential equations for slower learning, also called long term memory (LTM) or adaptive weight equations; and differential equations for processes on an intermediate time scale, such as habituation or desensitization. Just a few such equations, and their variations due to specialized parameter choices, form the foundation for most neural network models. This fact may in the future provide an opportunity for their systematic fabrication.

A larger number of design principles exist which guide the selection and relationships among these equations. This selection process amounts to choices of specialized parameters among the small number of general dynamical equations. Design principles help to identify modular families of models whose members all share aspects of one or more basic design ideas. These modules include objects called autoassociators, the LMS algorithm, back propagation (multi-layer LMS networks), instars, competitive learning schemes (competitively interacting instars), outstars, associative pattern learning networks (nonlinearly interacting arrays of outstars), associative mapping schemes (multi-layer arrays of hierarchical instar-outstar units in competition), avalanches (temporally organized arrays of outstars, instars, or reciprocal instar-outstar feedback networks), gated dipole opponent processors, additive or shunting cooperative-competitive feedack networks, gated dipole fields (arrays of gated dipoles linked by cooperative-competitive feedback networks), nonlinearly gated diffusive filling-in networks, and so on.

Recently, a number of journals have published special issues about neural networks that include many recent research developments and applications of such modules. These include the December 1, 1987 issue of *Applied Optics*; the January 1988, April 1988, and subsequent issues of *Neural Networks*; and the March 1988 issue of *Computer* magazine. Some books published within the past two years with illustrative applications are the two PDP books edited by James McClelland and David Rumelhart (MIT Press, 1986), the book edited by John Denker called **Neural Networks for Computing** (American Institute of Physics, 1986), and the books that I edited called **The Adaptive Brain**

(Elsevier/North-Holland, 1987) and **Neural Networks and Natural Intelligence** (MIT Press, 1988). A remarkably extensive work, containing approximately 500 contributions, is the four-volume conference proceedings edited by Charles Butler and Maureen Caudill for the IEEE meeting on neural networks that was held in San Diego in June, 1987 (order from IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854). Then there is the July 1987 NASA Survey of Artificial Neural Systems, edited by Dan Greenwood, that abstracts hundreds of recent contributions to the field.

## THE HISTORICAL ORDERING OF DESIGN IDEAS

One of the most striking sociological facts about the recent progress of neural network research is that new practitioners of the field have tended to rediscover these modules in the approximate order in which they were originally discovered and developed over the past few decades. In the last five years, for example, a new wave of interest has focussed first on autoassociator models, then LMS and multi-level LMS models, then on competitive learning models, and now is rapidly applying variations of the avalanche model. The most interesting aspect of this for me personally is that the modules among this list that I helped to pioneer, such as autoassociators, competitive learning, and avalanches— along with colleagues such as Shun-ichi Amari, James Anderson, Leon Cooper, Kunihiko Fukushima, Teuvo Kohonen, and Christoph von der Malsburg—were often found hard to read, obscure, and generally difficult to understand when they were first dicovered. Now that social conditions are right for assimilating the design intuitions that went into these discoveries, they seem so self-evident and elementary that many people find it hard to understand what a major intellectual struggle went into their original discovery. Even the use of nonlinear differential equations to express such computational ideas was once regarded with bafflement, derision, and scorn.

## PICTURE THE EMERGENT DYNAMICS

This fact may be helpful to engineers who may at first find it hard to read some of the neural network research that represents the cutting edge of advanced design ideas. As before, the problem now is not with writing style or even mathematical technique. The problem is to intuitively grasp the novel design concepts that underly a new module or architecture. Such design ideas are, at bottom, what make neural networks unique. People would not have any difficulty reading advanced neural network research if it was simply a way of packaging old ideas in new network realizations.

A good part of the difficulty has always derived from the key fact that "the architecture is the algorithm." One cannot simply read a list of rules to understand a neural network. The list of rules for the network is just a handful of differential equations. Although these equations provide a complete *formal* description of the network, they do not provide a *functional* understanding of the network, because the functional units which govern the architecture's problem-solving competence are *emergent properties* due to nonlinear interactions across the network.

To intuitively understand such a network one thus needs to grasp the relationship between the network's formal description and its emergent functional properties. I have been picturing the vibrant emergent dynamics of nonlinear neural networks for so long in my mind that I am occasionally shocked, especially when I am particularly tired and the picture momentarily fades, to realize how empty and meaningless such equations can seem, despite their formal rigor and completeness, to people who have not yet learned to think intuitively and pictorially about them.

# ARCHITECTURES FOR REAL-TIME INDIVIDUAL ADAPTIVE RESPONSE TO NONSTATIONARY ENVIRONMENTS

Neural network architectures typically join together several modules in a carefully crafted circuit. The most advanced architectures are aimed at achieving maximal *flexibility* and *autonomy* for providing general-purpose solutions to *modal problems*. I will illustrate what I mean by these remarks by sketching how my colleagues and I at the Center for Adaptive Systems (CAS) at Boston University go about developing such architectures.

The CAS carries out two types of theoretical activities which, although conceptually independent, have turned out to be mutually reinforcing in our case. One type of activity studies the fundamental design principles and mechanisms needed to explain and predict large data bases about brain and behavior. The other type of activity generates novel architectures for implementation as intelligent machines in technological applications. Why does the type of research done at CAS lend itself to both biological and technological applications? This research has proved to be relevant to technology both because of the types of problems we study and the methods that we use to solve them.

In particular, we study problems requiring real-time adaptive responses of individuals to unexpected changes in complex environments. These are the types of problems that humans and mammals need to solve in order to survive. These are also among the types of technological problems that traditional scientific and engineering approaches have not already well-handled.

Our methods for attacking such problems are systematic and rigorous. We typically begin by analysing a huge interdisciplinary data base concerning brain and behavior within a prescribed problem area. In our work on developing a neural network architecture for preattentive vision, for example, we have studied data from many parts of the vision literature—data about boundary completion, texture segmentation, surface perception, depth perception, motion perception, illusory figures, stabilized images, hyperacuity, brightness and color paradoxes, multiple scale filtering, and neurophysiology and anatomy from retina to prestriate cortex. Only through the sustained analysis of many hundreds or even thousands of such experiments can one accumulate enough data constraints to discard superficial modelling ideas and to discern a small number of fundamental design principles and circuits.

Such concepts do not make themselves known however, through a purely bottom-up shifting among huge heaps of data. They come into view by thinking about how these data could arise as emergent, or interactive, properties of a real-time process engaged moment-by-moment by the external visual environment. Being able to think in real-time about an immense mass of static data is an art. It is, I believe, the rate-limiting skill in this sort of work. Thus these real-time processes gradually become discernable through the active confrontation of a huge data mass with known theoretical principles, mechanisms, and computations about real-time neural network processes to test for matches and mismatches. Through this approach, a series of design paradoxes, or trade-offs, come into view which balance many data and computational requirements against one another. Gradually the accumulated impact of these design tradeoffs creates such an intense intellectual pressure within the emerging scheme of ideas that every fact and hypothesis ramifies through it with multiple implications. By this point, a well-defined family of real-time neural network architectures has usually come into view, supported by a new computational theory that is often quite invisible to a merely passive compilation of the data.

We develop and test these architectures using rigorous mathematical techniques and systematic parametric series of computer simulations in order to gain a complete formal understanding of their emergent, or interactive, properties. The combination of working on problems for which both biology and technology need answers, and developing these answers into rigorously characterized computational structures, makes such work equally

5

applicable to quantitative data analysis and to efficient technology transfer.

This rigorous approach has led to real-time neural network architectures that provide explicit examples of intelligent systems which overcome classical bottlenecks in stability, adaptability, scalability, capacity, and speed that have hampered the further development of AI algorithms. The demonstrations which guarantee these properties take the form of rigorous mathematical theorems and parametric computational analyses which have provided a firm foundation upon which software and hardware applications may confidently be supported.

## GENERAL PURPOSE SOLUTIONS TO MODAL PROBLEMS

What kind of problem can such a neural network architecture solve? Each architecture is being developed to supply a general-purpose solution within a focussed problem domain—what has been called a solution of a *modal problem*. A modal architecture is less general than a general-purpose digital computer but much more general than a typical AI algorithm.

The areas in which such modal neural network architectures are being developed include:

**perception**, notably innovations in biological and machine vision and speech, with applications to multidimensional image filtering, fusion, segmentation, completion, articulatory-to-auditory priming, automatically gain-controlled working memory, and self-scaling adaptive coding;

**cognitive information processing**, including new architectures for invariant adaptive pattern recognition, nonstationary hypothesis testing, self-adjusting parallel memory search, distributed decision-making under risk, and automatic reallocation of attention resources;

**cognitive-emotional interactions**, including architectures for rapidly focussing attention on environmental events and hypotheses which predict behavioral success based upon prior satisfaction of internal constraints, as in the action of rewards, punishments, homeostatic rhythms, or the unexpected nonoccurrence of expected goals; and

**goal-oriented motor control and robotics**, including architectures which circumvent classical combinatorial explosions to show how invariant properties of flexible eye and arm trajectories can be generated as emergent real-time properties of nonlinear neural interactions, rather than as explicitly pre-planned commands, and how self-calibration of movement command parameters can be learned automatically after partial accidents or other unexpected environmental feedback.

Although each of these projects can, at least in part, be carried out independently, they can also collectively benefit from efficiencies of cooperation, interfacing, and scale if they are organized as part of a coordinated research program aimed at the design of intelligent machines capable of autonomous adaptive real-time operation in unanticipated environmental situations. Interdisciplinary institutes such as CAS provide the type of research core around which such larger coordinated projects can effectively be planned. More of this type of coordination may very well be a next step in the institutional development of our field.

## A NEURAL NETWORK ARCHITECTURE
## FOR PREATTENTIVE VISION

My colleagues Michael Cohen, Ennio Mingolla, Dejan Todorović, and I have, for example, been developing a neural network architecture for general-purpose preattentive vision. Many AI algorithms for machine vision have been too specialized for applications to real-world problems. Such algorithms are often designed to deal with one type

of information—for example, boundary, disparity, curvature, shading, or spatial frequency information. Moreover, such algorithms typically use different computational schemes to analyse each distinct type of information, so that unification into a single general-purpose vision algorithm is difficult at best. For such AI algorithms, other types of signals are often contaminants, or noise elements, rather than cooperative sources of ambiguity-reducing information. Unfortunately, most realistic scenes contain partial information of several different types in each part of a scene.

In contrast, when we humans gaze upon a scene, our brains rapidly combine several different types of locally ambiguous visual information to generate a globally consistent and unambiguous representation of form-and-color-in-depth. This state of affairs raises the question: What new computational principles and mechanisms are needed to understand how multiple sources of visual information cooperate automatically to generate a percept of 3-dimensional form?
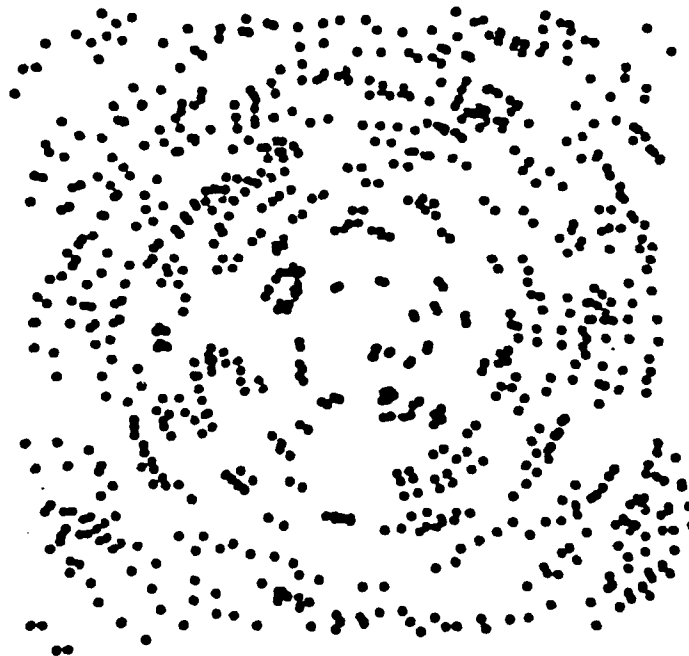


**Figure 1.** A Glass pattern: The emergent circular pattern is "recognized," although it is not "seen," as a pattern of differing contrasts.

The CAS vision architecture clarifies how scenic data about boundaries, textures, shading, depth, multiple spatial scales, and motion can be cooperatively synthesized in real-time into a coherently fused representation of 3-dimensional form. Moreover, it has become clear through collaborative work with colleagues at M.I.T. Lincoln Laboratory that the same processes which are useful to automatically process visual data from human sensors are equally valuable for processing noisy multidimensional data from artificial sensors, such as laser radars. These processes are called emergent segmentation and featural filling-in. The inability of alternative theories to mechanistically understand these processes has been a major bottleneck in the development of quantitative visual theory and computer vision applications.
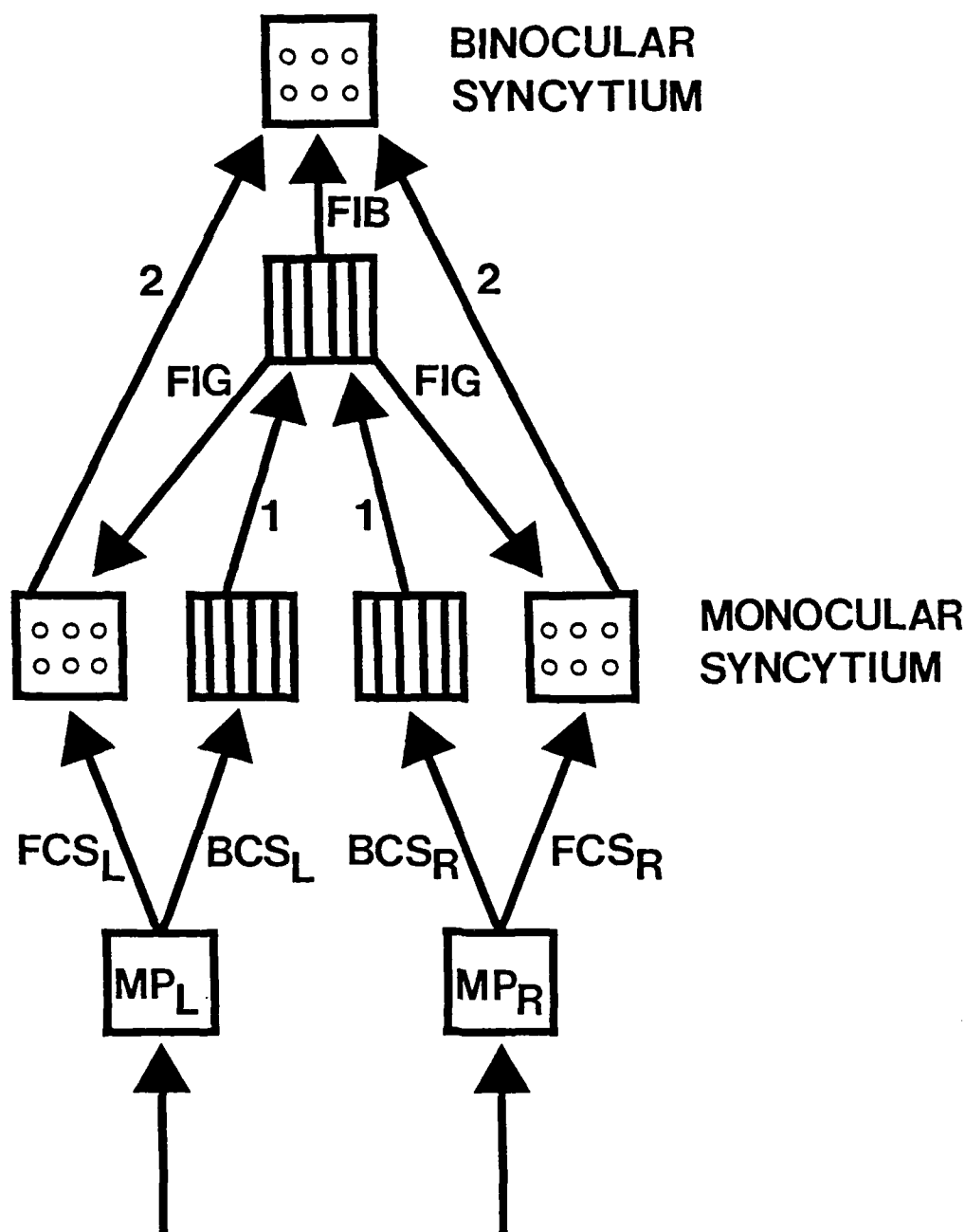
7

**Figure 2.** Macrocircuit of monocular and binocular interactions within the Boundary Contour System (BCS) and the Feature Contour System (FCS): Left and right monocular preprocessing stages ($MP_L$ and $MP_R$) send parallel monocular inputs to the BCS (boxes with vertical lines) and the FCS (boxes with three pairs of circles). The monocular $BCS_L$ and $BCS_R$ interact via bottom-up pathways labelled 1 to generate a coherent binocular boundary segmentation. This segmentation generates output signals called filling-in generators (FIGs) and filling-in barriers (FIBs). The FIGs input to the monocular syncytia of the FCS. The FIBs input to the binocular syncytia of the FCS. Inputs from the MP stages interact with FIGs at the monocular syncytia to selectively generate binocularly consistent Feature Contour signals along the pathways labelled 2 to the binocular syncytia. These monocular Feature Contour signals interact with FIB signals to generate a multiple scale representation of form-and-color-in-depth within the binocular syncytia.

8

**Figure 3.** Circuit diagram of the Boundary Contour System: Inputs activate oriented masks of opposite direction-of-contrast which cooperate at each position and orientation before feeding into an on-center off-surround interaction. This interaction excites like-orientations at the same position and inhibits like-orientations at nearby positions. The affected cells are on-cells within a gated dipole field. On-cells at a fixed position compete among orientations. On-cells also inhibit off-cells which represent the same position and orientation. Off-cells at each position, in turn, compete among orientations. Both on-cells and off-cells are tonically active. Net excitation of an on-cell excites a similarly oriented cooperative receptive field of a bipole cell at a location corresponding to that of the on-cell. Net excitation of an off-cell inhibits a similarly oriented cooperative receptive field of a bipole cell at a location corresponding to that of the off-cell. Thus, bottom-up excitation of a vertical on-cell, by inhibiting the horizontal on-cell at that position, disinhibits the horizontal off-cell at that position, which in turn inhibits (almost) horizontally oriented cooperative receptive fields that include its position. Sufficiently strong net positive activation of both receptive fields of a cooperative cell enables it to generate feedback via an on-center off-surround interaction among like-oriented cells. On-cells which receive the most favorable combination of bottom-up signals and top-down signals generate the emergent perceptual grouping.

9

The difficulties inherent in computationally understanding biological vision can be appreciated by considering the following example. Consider the visual image that is depicted below, called a Glass pattern (Figure 1). When we view such a Glass pattern, we *see* many black dots on white paper, but we also *recognize* circular groupings, which are often called *emergent segmentations*. For most individuals, these circular groupings do not generate brightnesses or colors that differ significantly from the background. Thus there is a profound difference between seeing and recognizing, and we can sometimes recognize groupings that we cannot see. This state of affairs raises the interesting question: If we can recognize things that we cannot see, then why do we bother to see?

In order to computationally understand such labile relationships between recognized emergent segmentations and seen brightnesses, the theory shows that there exist fundamental limitations of the visual measurement process—that is, newly discovered uncertainty principles are just as important in vision as they were in providing a foundation for quantum mechanics. The theory analyses how the nervous system as a whole can compensate for these uncertainties using both parallel and hierarchical stages of neural processing. Thus the nervous system is designed to achieve *heterarchical compensation for uncertainties of measurement*, and these particular compensations lead to qualitatively new designs of automatic vision machines.

Figure 2 provides a macrocircuit of the CAS vision architecture. Each box in the figure includes one or more layered nonlinear neural networks. Figure 3 schematizes one of the major subsystems of the theory, called the Boundary Contour System, which automatically generates a form-sensitive emergent boundary segmentation of a scene. The BC System operates autonomously using nonparametric, internally regulated nonlinear feedback loops. It is thus quite unlike stochastic relaxation techniques, such as simulated annealing, which rely on the independent, external manipulation of a noise parameter and predetermined probability distributions to regulate convergence to equilibrium. Consequently, stochastic relaxation techniques can sharpen expected properties of an image, but they are unable to simulate a key property needed for designing a general-purpose vision preprocessor: The automatic discovery of emergent image groupings that may never have been experienced before, such as the circular groupings among the dots of the Glass pattern in Figure 1.

## SELF-ORGANIZING PATTERN CLASSIFIER: ART 1 AND 2

Output patterns from an autonomous vision preprocessor get fed as inputs into an autonomous self-organizing pattern classifier, called an Adaptive Resonance Theory (or ART) circuit. I originally introduced ART in 1976 as a cognitive theory aimed at answering the following basic question about autonomous behavior, called the stability-plasticity dilemma.

How can a learning system be designed to remain plastic, or adaptive, in response to significant events, yet also remain stable in response to irrelevant events? How does the system know how to switch between its stable and its plastic modes to achieve stability without rigidity and plasticity without chaos? In particular, how do we continue to learn new things without being forced to forget everything we ever learned before? Moreover, how does a system accomplish this balancing act without using a teacher? Thus, Adaptive Resonance Theory was introduced to help explain how humans and animals can learn, on their own, to cope so well with a world of seemingly endless richness and complexity whose rules can change unexpectedly.

The theory has provided mechanistically precise answers to such fundamental questions as:

Why do we pay attention? Why do we learn expectations about the world? In particular, how do we cope so well with unexpected events, and how do we manage to do so as well as we do when we are on our own, and do not have a teacher as a guide? How do we spontaneously discover, test, and learn hypotheses about an ever-changing world? How do
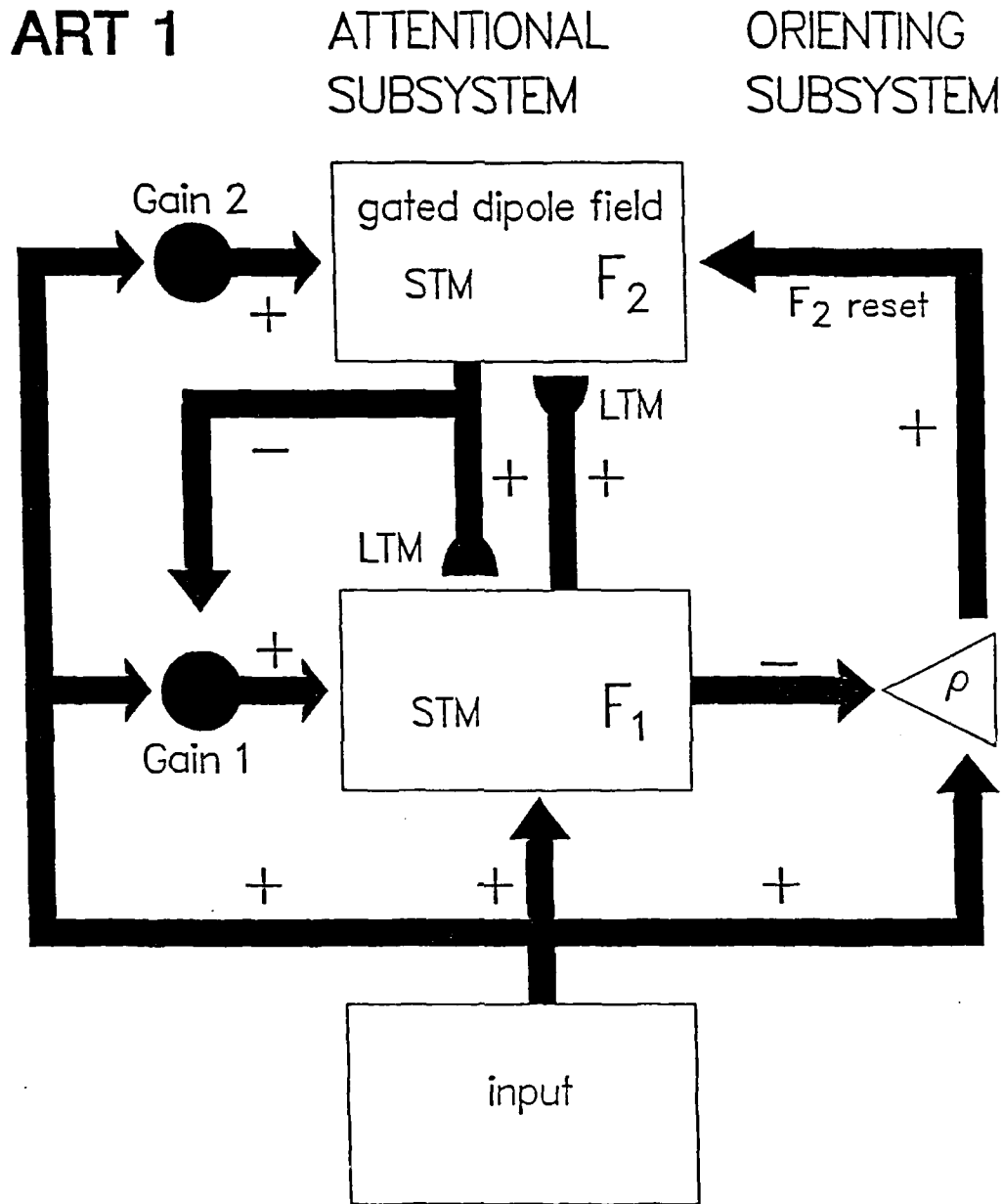
**Figure 4.** A typical ART 1 architecture. Rectangles represent fields where STM patterns are stored. Semicircles represent adaptive filter pathways and arrows represent paths which are not adaptive. Filled circles represent gain control nuclei, which sum input signals. Their output paths are nonspecific in the sense that at any given time a uniform signal is sent to all nodes in a receptor field. Gain control at $F_1$ and $F_2$ coordinates STM processing with input presentation rate. The orienting subsystem generates a reset wave to $F_2$ when sufficiently large mismatches between bottom-up and top-down patterns occur at $F_1$. This reset wave selectively and enduringly inhibits previously active $F_2$ cells until the input is shut off, and triggers an automatic hypothesis testing cycle that searches for an appropriate code for the input pattern.
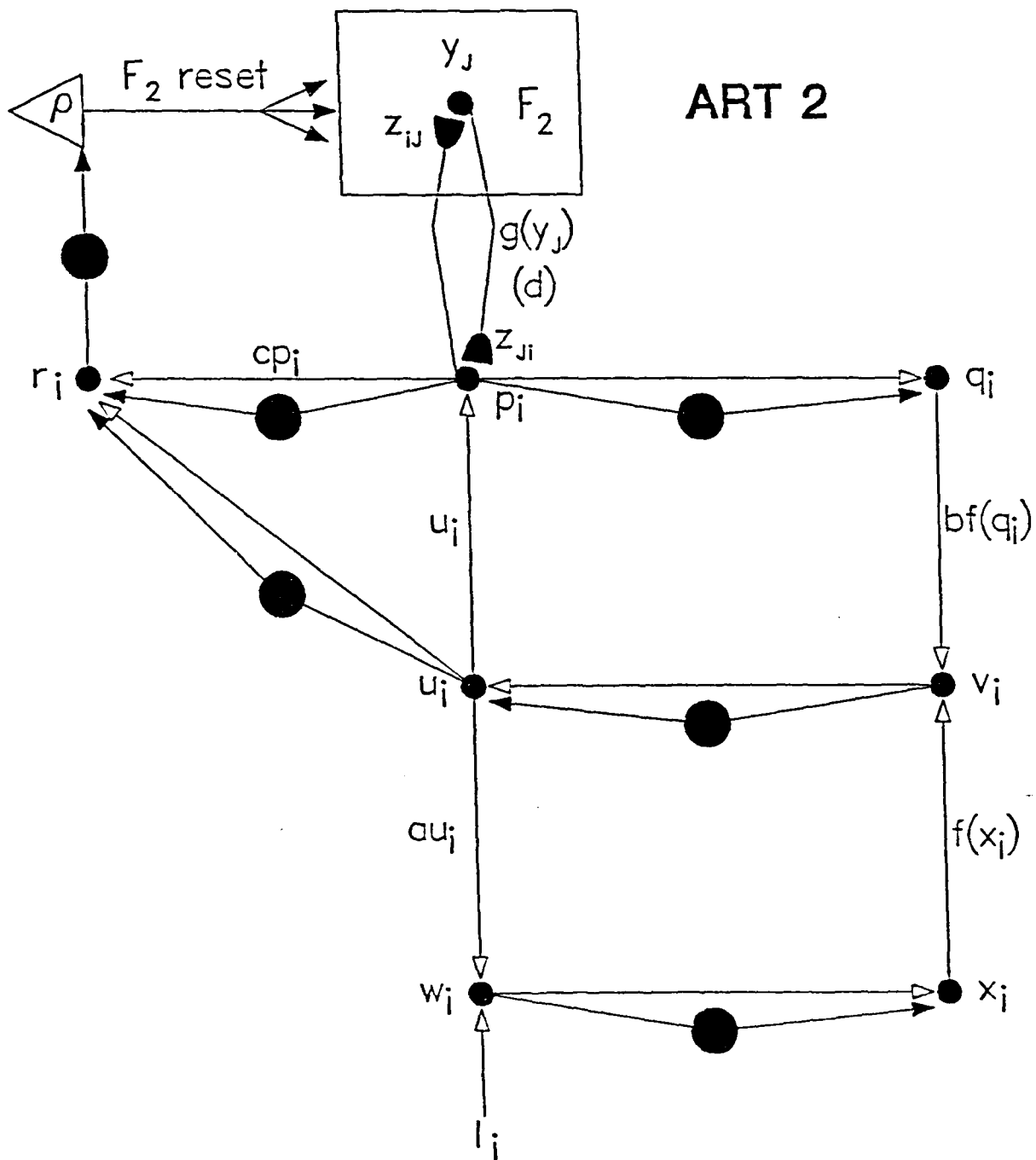
**Figure 5.** A typical ART 2 architecture. Open arrows indicate specific patterned inputs to target nodes. Filled arrows indicate nonspecific gain control inputs. The gain control nuclei (large filled circles) nonspecifically inhibit target nodes in proportion to the $L_2$-norm of STM activity in their source fields. As in ART 1, gain control (not shown) coordinates STM processing with input presentation rate. The hemi-disks signify the location of adaptive LTM traces.

# TABLE 1

| ART Architecture | Alternative Learning Properties |
|---|---|
| Real-time (on-line) learning | Lab-time (off-line) learning |
| Nonstationary world | Stationary world |
| Self-organizing (unsupervised) | Teacher supplies correct answer (supervised) |
| Memory self-stabilizes in response to arbitrarily many inputs | Capacity catastrophe in response to arbitrarily many inputs |
| Effective use of full memory capacity | Can only use partial memory capacity |
| Maintain plasticity in an unexpected world | Externally shut off plasticity to prevent capacity catastrophe |
| Learn internal top-down expectations | Externally impose costs |
| Active attentional focus regulates learning | Passive learning |
| Slow or fast learning | Slow learning or oscillation catastrophe |
| Learn in approximate-match phase | Learn in mismatch phase |
| Use self-regulating hypothesis testing to globally reorganize the energy landscape | Use noise to perturb system out of local minima in a fixed energy landscape |
| Fast adaptive search for best match | Search tree |
| Rapid direct access to codes of familiar events | Recognition time increases with code complexity |
| Variable error criterion (vigilance parameter) sets coarseness of recognition code in response to environmental feedback | Fixed error criterion in response to environmental feedback |
| All properties scale to arbitrarily large system capacities | Key properties deteriorate as system capacity is increased |

we know what combinations of facts are useful for dealing with a given situation and what combinations of facts are irrelevant? How do we recognize familiar facts so quickly even though we may know many other things? How do we join together knowledge about the external world with information about our internal needs to quickly make decisions that have a good chance of satisfying these needs? Finally, what do all of these properties have in common?

My colleague Gail Carpenter and I have developed two generations of ART architectures, called ART 1 and ART 2. ART 1 (Figure 4) is a neural network architecture that self-organizes a stable pattern recognition code in response to arbitrary sequences of *binary* input patterns. ART 2 (Figure 5) self-organizes stable pattern recognition codes in response to arbitrary sequences of *analog* (including binary) input patterns. The guarantee of being able to learn a pattern recognition code in response to *arbitrary* input sequences shows that these ART systems provide a general purpose solution to a modal problem, and opens the possibility of using them or their decendents in autonomous machines which may safely be confronted by an unexpected, nonstationary pattern sequence while on the job. Table 1 outlines some of the differences between properties of ART architectures and a variety of alternative approaches to pattern recognition, such as back propagation, simulated annealing, or rule-based systems.

Within such an ART architecture, the process of adaptive pattern recognition is a special case of the more general cognitive process of hypothesis discovery, testing, search, classification, and learning. This latter property opens the possibility of applying ART systems to more general problems of adaptively processing large abstract information sources and data bases.

These vision preprocessor and ART autonomous classifier examples are just two of the many neural network architectures now being developed by engineers and scientists worldwide. Some of them provide a fertile ground for gaining a new understanding of biological intelligence. Others suggest novel computational theories with natural realizations as real-time adaptive neural network architectures with promising properties for tackling some of the outstanding problems in computer science and technology today. Still others do both. Whatever the focus, here is a field ready to challenge and reward the sustained efforts of a wide variety of gifted people.

END

DATE
FILMED
6-1988
DTIC